Using Fuzzy Clustering for Real-time Space Flight Safety

Charles Lee
Science Applications International Corporation
NASA Ames Research Center
Moffett Field, CA. 94035
650-604-6054
clee@mail.arc.nasa.gov

Richard E. Haskell
Department of Computer Science and Engineering
Oakland University
Rochester, MI. 48092
248-370-2861
haskell@oakland.edu

Darrin M. Hanna
Department of Computer Science and Engineering
Oakland University
Rochester, MI. 48092
248-370-2170
dmhanna@oakland.edu

Richard L. Alena
NASA Ames Research Center
Moffett Field, CA. 94035
650-604-0262
Richard.L.Alena@nasa.gov

Abstract

To ensure space flight safety, it is necessary to monitor myriad sensor readings on the ground and in flight. Since a space shuttle has many sensors, monitoring data and drawing conclusions from information contained within the data in real time is challenging. The nature of the information can be critical to the success of the mission and safety of the crew and therefore, must be processed with minimal data-processing time. Data analysis algorithms could be used to synthesize sensor readings and compare data associated with normal operation with the data obtained that contain fault patterns to draw conclusions. Detecting abnormal operation during early stages in the transition from safe to unsafe operation requires a large amount of historical data that can be categorized into different classes (non-risk, risk). Even though the more than 20 years of shuttle flight program has accumulated volumes of historical data, these data don't comprehensively represent all possible fault patterns since fault patterns are usually unknown before the fault occurs. This paper presents a method that uses a similarity measure between fuzzy clusters to detect possible faults in real time. A clustering technique based on a fuzzy equivalence relation is used to characterize temporal data. Data collected during an initial time period are separated into clusters. These clusters are characterized by their centroids. Clusters formed during subsequent time periods are either merged with an existing cluster or added to the cluster list. The resulting list of cluster centroids, called a cluster group, characterizes the behavior of a particular set of temporal data. The degree to which new clusters formed in a subsequent time period are similar to the cluster group is characterized by a similarity measure, q. This method is applied to downlink data from Columbia flights. The results show that this technique can detect an unexpected fault that has not been present in the training data set.

1. Introduction

Space flight safety highly relies on the fault detection at very early stages of the normal to fault transition period. Monitoring the sensor readings on the ground and in flight is one of the very important measures to fulfill such a task. This paper presents one method to detect the faults by using a fuzzy clustering algorithm.

Many different clustering techniques have been used for analyzing multivariate data [1]. These methods have been applied to problems in knowledge discovery and data mining [2,15]. The standard clustering algorithms assign each data sample to one of many clusters in which all samples in a particular cluster are similar in some sense. Fuzzy clustering algorithms do not insist that each sample must belong to only one cluster, but rather samples can belong to more than one cluster to varying degrees. The most well known fuzzy clustering algorithm is the fuzzy c-means algorithm [3] that requires that the number of cluster centers, c, be given. A different clustering approach that does not require the number of clusters to be known beforehand is based on the use of fuzzy equivalence relations [4, 5]. In this method a fuzzy compatibility relation matrix, Q, is formed in which each entry in the matrix represents the degree to which two different samples are close to each other. A value of 1 (on the main diagonal) represents the degree to which a sample is close to itself, while a value of 0 represents samples separated by the largest possible distance in the data set. The transitive closure of Q will induce crisp partitions of the data (resulting in different numbers of clusters) by choosing different α-cuts of a fuzzy set [5]. Clusters formed in this manner will be used in this paper to characterize pattern behavior over time of space flight data.

Fuzzy clustering can not only extract features from raw data directly [6], but also select the optimal feature sets or reduce the dimensionality of obtained features [7]. Fuzzy clustering is also playing an important role in the raw data domain as well as in transform domains. While most of the existing fuzzy clustering algorithms are proposed for static data sets, it also needs to be studied in the case of dynamic data sets. This paper involves dynamic fuzzy data sets. It is impossible for a single clustering criterion to solve all the possible problems of unsupervised classification. There are many functions of similarity, such as maximum likelihood function [8], maximum entropy criterion [9], minimum volume criterion [10] and information criterion [11], and other approaches [14,15]. This paper will use a cluster similarity measure to detect the onset of a risk situation in space flight data. A similar approach has been used to detect driver behavior patterns [12].

Assume that multi-variable space flight data are collected for an initial amount of time, Δt_1. These data are clustered to form c_1 clusters. During the next interval of time, Δt_2, new data are collected and clustered to form c_2 clusters. If the sensor activity during Δt_2 is similar to the sensor activity during Δt_1 then the c_2 clusters (as specified by their centroids) will be similar to the c_1 clusters. On the other hand, if the sensor activity has changed somewhat, some of the c_2 clusters will be different from the c_1 clusters. Clusters in c_2 that are close to clusters in c_1 will be merged with the corresponding clusters in c_1. The remaining clusters in c_2 will be added to the c_1 clusters to form a new, larger c_1 cluster set. This process is continued for each successive interval of time, Δt_n, with new clusters either being merged with an existing cluster or added to the cluster set. After a while, the resulting group of clusters, called a *cluster group*, will characterize a particular space flight pattern. This method will allow a large amount of data collected over a period of time to be clustered in a manner that uses a manageable amount of data at each clustering step.

Once a cluster group is formed, new data that are collected during a subsequent interval of time, Δt_n, will form its own cluster set. In this paper we will use a similarity measure, q, that will measure the degree to which this new cluster set "fits in" with the cluster group. By using this similarity measure one can determine which of several cluster groups provides the best fit for a particular cluster set. Once the cluster group has stabilized for a particular space flight and a new cluster appears that is significantly different from what has been seen before, this may indicate some type of risk situation that would require immediate attention.

2. Clustering Based on a Fuzzy Equivalence Relation

This section will describe a clustering technique in which multivariate data are used to form a fuzzy equivalence relation matrix. Different α-cuts of this fuzzy set will produce a different number of clusters of the original data. Given a data set $\{(x_{11},\ldots,x_{1p}),\ldots,(x_{n1},\ldots,x_{np})\}$ with n

$$M_k = \max_{j\in\{1,..,n\}} x_{jk}$$

samples over a p-dimensional feature space, P, a fuzzy compatibility relation matrix Q with

$$m_k = \min_{j\in\{1,..,n\}} x_{jk}$$

dimension $n \times n$ is computed. Define M_k and m_k as the maximum and minimum data point x_{jk} for each feature k in P as

and

Define the i,j-th entry in Q as

to form the fuzzy compatibility relation matrix Q. Each q_{ij} represents the degree to which data point x_i is close to data point x_j. The distance measure

$$q_{ij} = 1 - \frac{1}{p}\sum_{k=1}^{p}\left(\left|\frac{x_{ik} - x_{jk}}{M_k - m_k}\right|^s\right)^{1/s} \qquad (1)$$

in Eq. (1) will be the Hamming distance for $s = 1$ and the Euclidean distance for $s = 2$. In our experiments we will use a value of $s = 1$.

The matrix Q is symmetric and reflexive. However, the generalization of transitivity to fuzzy relations is not unique [5]. One common

$$Q(x,z) \geq \max_{y\in Y} \min\left[Q(x,y), Q(y,z)\right]$$

definition is to say that a fuzzy relation Q is transitive if and only if

The right-hand side of this inequality represents the composition of relation Q with itself, $Q \circ Q$. The following algorithm can compute the transitive closure, T, of Q.

Transitive Closure:

$$
\begin{aligned}
&do\\
&\quad T' = Q;\\
&\quad Q = T' \circ T';\\
&while(Q \neq T')\\
&T = T'
\end{aligned}
$$

From the transitive closure matrix T with elements t_{ij}, a collection of clusters, C, is formed

$$\forall i, j \in C_k, t_{ij} \geq \alpha$$

for a specific membership degree α. Set $C_k \in C$ such that
forms a fuzzy equivalence class. Define a fuzzy equivalence cluster W_i by $\{(x_{w1},\ldots,x_{wP}) \mid w \in C_i\}$.

Multiple Cluster Sets

Define the centroid, a_i, of cluster W_i by

$$a_i = \frac{1}{|C_i|} \sum_{j \in C_i} x_j$$

where $|C_i|$ denotes the cardinality of cluster set C_i. Let $A_\alpha = \{a_1, \ldots, a_m\}$ be the set of fuzzy cluster centroids resulting from data collected during a time interval, Δt_{k1} and $B_\alpha = \{b_1, \ldots, b_n\}$ be a second set of fuzzy cluster centroids resulting

$$\max_{Ak} = \max_i \max_{j \in C_i} x_{jk}$$

from data collected during a time interval, Δt_{k2} where $k2 > k1$ with accumulative maximums and minimums for each feature, max_{Ak}, max_{Bk}, min_{Ak}, and min_{Bk}, where, without loss of generality,
and

$$\min_{Ak} = \min_i \min_{j \in C_i} x_{jk}$$

and weight vectors $W_{a\alpha} = \{w_{a1}, \ldots, w_{am}\}$ and $W_{b\alpha} = \{w_{b1}, \ldots, w_{bn}\}$ given, without loss of generality, by

$$W_{a\alpha} = \{w_{ai} \mid w_{ai} = |C_i|\}$$

Define r_k as the global range,

$$r_k = \max(\max_{Ak}, \max_{Bk}) - \min(\min_{Ak}, \min_{Bk})$$

Form the fuzzy relation matrix, Z, by

$$z_{ij} = 1 - \frac{1}{p} \sum_{k=1}^{p} \left(\left| \frac{(a_{jk} - b_{ik})}{r_k} \right|^s \right)^{1/s} \quad (2)$$

where p is the dimension of the feature space and $s = 1$ for a Hamming distance or $s = 2$ for a Euclidean distance. Form the projections

$$\rho_i^A = \max_i z_{ij} \quad (3)$$

and

$$\rho_i^B = \max_j z_{ij} \quad (4)$$

Merging Cluster Sets

Using the centroid relation matrix, Z, and ρ_i^B, a new collection of clusters $\aleph_{t+1}$ is constructed with threshold, β. For all $\rho_i^B > \beta$, a_j is replaced as follows,

$$\left. \frac{w_{aj} a_j + w_{bi} b_i}{w_{aj} + w_{bi}} \right|_{z_{ij} = \rho_i^B} \quad (5)$$

Accordingly, w_{aj} is updated with $w_{bi} + w_{aj}$.

Finally, form

$$\aleph_{t+1} = \left\{ i \mid \rho_i^B \leq \beta \right\} \cup A_\alpha \quad (6)$$

for a new collection of clusters representing time-series Δt_{k1} and Δt_{k2}. This method is repeated for each successive time interval, Δt.

Cluster Similarity Measure

Let A_α be the set of clusters formed by adding and merging cluster sets over a number of time intervals. Let B_α be the cluster set during a new time interval, Δt_n. A fuzzy relation matrix Z can be computed from Eq. (2) and the projection ρ_i^B is given by Eq. (4). The similarity measure q, is defined as the degree to which cluster set B_α is similar to cluster set A_α as

$$q = \frac{1}{n} \sum_{i=1}^{n} \rho_i^B \qquad (7)$$

This clustering method based on a fuzzy equivalence relation will be used in the following section to detect the onset of a risk situation in space flight data.

3. Test Results

The space shuttle program has accumulated a significant amount of data from past flights. To detect an abnormal status, we used Columbia shuttle data as input. The data are 18 sensor readings from the left wing. To test the fuzzy clustering method described in Section 2 for fault detection on the shuttle flight, we selected sample data from space shuttle STS107 that included the time period where we expected the onset of an abnormal status. Our goal of the experiments is to determine whether the algorithm properly detect new cluster in the time interval of the failure. The data samples were divided into six separate files labeled 1 – 6 in increasing time sequence. The results of the clustering experiment are summarized in Table 1.

Table 1 Results of clustering space flight data (α = 0.8, β = 0.8)

File no.	No. of samples	No. of clusters	Cluster	Merge	total no.of cluster	No. of samples in clusters
1	100	1	1		1	100
2	100	1	2	(1,2)	1	200
3	100	1	3	(1,2,3)	1	300
4	100	1	4	(1,2,3), 4	2	300/100
5	88	1	5	(1,2,3), (4,5)	2	300/188
6	29	1	6	(1,2,3), (4,5,6)	2	300/217

File number 1 contained 100 samples. Using these samples a 100 x 100 fuzzy compatibility relation matrix Q is computed using Eq. (1). A corresponding fuzzy equivalence relation is found by computing the transitive closure of Q using the algorithm described in Section 2. An alpha cut, α = 0.8, results in a single cluster as indicated in Table 1.

The 100 samples in file number 2 are processed in the same way and also result in a single cluster with a somewhat different centroid from cluster 1 from file number 1. The merging operation described by Eqs. (5) and (6) is then carried out with a value of β = 0.8. In this case clusters 1 and 2 are similar enough to be merged as indicated in the second row of Table 1.

The same clustering and merging operations are then applied to the 100 samples in file number 3. These 100 samples form a single cluster (3) that is merged with the existing cluster (1,2) to form a single merged cluster (1,2,3).

When the same clustering and merging operations (with α = 0.8 and β = 0.8) are applied to the 100 samples in file number 4 a single cluster (4) is created, but it is not similar enough to the existing cluster (1,2,3) to be merged. It therefore starts a new cluster labeled 4.

Files number 5 and 6 each form a single cluster that are merged to the 4 cluster to eventually create the second cluster (4,5,6) as indicated in the last row of Table 1.

When the total data set was reviewed after creating these clusters it was found that the creating of the second merged cluster in file number 4 corresponded to the first indication of trouble in the space flight data, which is the time when one large piece and at least two small pieces of insulating foams separated from external tank left bipod ramp area [16]. Table 2 shows the time of those events.

Table 2. The time wheninsulating foams separated from the External Tank

Time from Launch	0	81.7second	81.9second
Events	Launch	1 large, 2 small foam separation	Foam Struck left wing

This suggests that this clustering method could be useful for the early detection of potential problems. This is 17 days before the accident happened when the Columbia space shuttle re-entry. By clustering and merging temporal multi-dimensional data in real time, the creation of new clusters may indicate a significant change in nominal operating conditions.

4. Conclusion

In this paper we have shown how clustering based on a fuzzy equivalence relation can be used to detect changes from normal operating conditions in space flight data. One advantage of the method is that large amounts of data that are arriving in real time can be absorbed by clustering small segments of data and merging the clusters into cluster groups that characterize a particular time sequence. A similarity measure, q, measures how closely a new cluster set matches (fits into) a larger cluster group. This can be used to determine if the new cluster set belongs to a particular cluster group. The individual components of the fuzzy projection used to calculate q can be used to indicate when a new cluster that is different from existing clusters appears unexpectedly. This could indicate a potential problem associated with the space flight data.

REFERENCES

[1] B. S. Everitt, *Cluster Analysis*, 3rd Ed., Edward Arnold, London, 1993.

[2] D. Wishart, "Efficient hierarchical cluster analysis for data mining and knowledge discovery," Computer Science and Statistics, 30, pp. 257-263, 1998.

[3] J. D. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[4] G. J. Klir, U. St.Clair and B. Yuan, *Fuzzy Set Theory – Foundations and Applications*, Prentice Hall PTR, Upper Saddle River, NJ, 1997.

[5] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic – Theory and Applications*, Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[6] V. Ramdas, V. Sridhar, G. Krishna, "An effective technique for feature extraction," Pattern Recognition Letters, 1994, 15:885.

[7] J. C. Bezdek, P. F. Castelaz, Prototype classification and feature selection with fuzzy sets, IEEE SMC, 1977, 2(7): 87.

[8] E. Trauwaert, L. Kaufman, P. Rousseeuw, Fuzzy clustering algorithms based on the maximum likelihood principle,Fuzzy Sets and System, 1991, 85(42): 213.

[9] R. P. Li, M. Mukaidino, A maximum entropy approach to fuzzy clustering, IEEE-FUZZ'95, 1995: 2227.

[10] J. C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well separated cluster, J. Cybernet,1974, 3: 32.

[11] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, New York: Plenum Press, 1981.

[12] R. E. Haskell, D. M. Hanna, P. Li, K. C. Cheok, and G. Hudas, "Finding Pattern Behavior in Temporal Data Using Fuzzy Clustering," Intelligent Enginering Systems Through Artificial Neural Networks, Vol. 10, Smart Engineering Design: Neural Networks, Fuzzy Logic, Evolutionary Programming, Data Mining, and Complex Systems, C. H. Dagli, et al., Eds. Proc. of the Artificial Neural Networks in Engineering Conference (ANNIE 2000), St. Louis, MO, pp. 703-711, Nov. 5-8, 2000.

[13] Y. L. Tseng, S. B. Yang, "A genetic approach to the automatic clustering problem", *Pattern Recognition*, Vol.34 (2), pp.415-424 (2001).

[14] N. B. Karayiannis, M. M. Randolph-Gips, Soft Learning Vector Quantization and Clustering Algorithms Based on Non-Euclidean Norms: Multinorm Algorithms, IEEE Transactions on Neural Networks, Vol. 14, No. 1, Jan. 2003

[15] D. M. Rocke and J. Dai, Sampling and Subsampling for Cluster Analysis in Data Mining:With Applications to Sky Survey Data, Data Mining and Knowledge Discovery, 7, 215–232, 2003

[16] Columbia Accident Investigation Board Report Volume I, August, 2003